

fragment of pBR322 is ligated to the backbone of LG1. The ligated DNA is used to transform competent E. coli that are plated on ampicillin-containing plates after a short grow-out.

5

Ampicillin-resistant colonies are picked. Plasmid DNA of the phagemid from the resistant colonies are digested with Bsu36I and RsrI. To verify the construction, DNA from phagemids with the correct restriction  
10 digestion pattern is sequenced: a) from about 20 bases above the Bsu36I site to about 20 bases below the RsrI site, and b) for about 30 bases either side of the unique AvaII site. The correct construct is named pLG2.

15 The AccI restriction site is no longer needed for vector construction. To eliminate this site, RF pLG2 dsDNA is cut with AccI, treated with Klenow fragment and dATP and dTTP to make it blunt and then religated. The ligated DNA is used to transform competent cells; after a  
20 short grow-out, ampicillin-resistant colonies are selected. Restriction digestion is used to screen phagemid DNA from these colonies; the desired product cannot be cut with AccI. To verify the construction, DNA from colonies lacking an AccI restriction site is  
25 sequenced from about 20 bases above the former AccI site to about 20 bases below it. The cloning vector, named pLG3, is now ready for stepwise insertion of the osp-ipbd gene.

30 We are now ready to design a gene (See Sec. 4) that will cause BPTI-domains to appear on the outer surface of an M13 derivative: LG7.

To obtain a novel protein domain attached to the  
35 outside of M13, we insert DNA that codes for mature BPTI

after A23 of the precoat protein of M13. Mature BPTI begins with an arginine residue, which is charged; cleavage by signal peptidase I is normal in such cases. Signal peptidase I (SP-I) cuts a chimera of M13 coat protein and BPTI after A23 leaving mature BPTI attached at its carboxy end to the amino terminus of M13 CP.

The following amino-acid sequence, called AA\_seq2, is constructed, by inserting the sequence for mature BPTI (shown underscored) immediately after the signal sequence of M13 precoat protein (indicated by the arrow) and before the sequence for the M13 CP.

AA\_seq2

15		1	1	2	↓ ↓ ↓ 2	3	3	4	4	5	
	5	0	5	0	5	0	5	0	5	0	
	<u>MKKSLVLKASVAVATLVPMLS</u> <u>FARPDFCLEPPYTGPC</u> <u>KARIIRYFYNAKA</u>										
20		5	6	6	7	7	8	8	9	9	10
		5	0	5	0	5	0	5	0	5	0
	<u>GLCOTFVYGGCRAKRNNFKSAEDCMRTC</u> <u>GGAEGDDPAKAAFNSLQASAT</u>										
25		10	11	11	12	12	13				
		5	0	5	0	5	0				
	<u>EYIGYAWAMVVVIVGATIGIKLFKKFTSKAS</u>										

30

We adopt the convention that sequence numbers of fusion proteins refer to the fusion, as coded, unless otherwise noted. Thus the alanine that begins M13 CP is referred to as "number 82", "number 1 of M13 CP", or "number 59 of the mature BPTI-M13 CP fusion".

The osp-ipbd gene is regulated by the lacUV5 promoter, so that the level of expression can be regulated by the concentration of IPTG supplied in the growth medium. (See Sec. 4.1). The host strain of E. coli

should harbor the lacI<sup>Q</sup> gene that represses the lacUV5 promoter to a greater extent than lacI<sup>+</sup>. The osp-ipbd gene is ended by the trp attenuator so that RNA polymerase will not read through into subsequent genes. The osp-ipbd gene is expressed and processed in parallel with the wild-type gene VIII. The novel protein, that consists of BPTI tethered to a M13 CP domain, constitutes only a fraction of the coat. Affinity separation is able to separate phage carrying only five or six copies of a molecule that has high affinity for an affinity matrix (SMIT85); 1% incorporation of the chimeric protein results in about 30 copies of the protein exposed on the surface. If this is insufficient, additional copies may be provided.

15

Figure 9 shows, in stereo, a hypothetical model of a short segment of the coat of a derivative of M13 in which some coat protein monomers are fusions of mature BPTI to the amino terminus of the normal M13 CP. The figure shows only protein  $\alpha$  helices; the DNA, not shown, lies inside the cylinder. The model of M13 coat is after the model for f1 of Marvin and colleagues (BANN81). The BPTI domain is taken from the Brookhaven Protein Data Bank entry "6PTI" and was attached by standard model building methods that insure that covalent bond lengths and angles are close to acceptable values. The space between the alpha helical main chains is filled by protein side groups so that the DNA is protected from solvent. The figure is not meant to suggest that BPTI fused to M13 CP will adopt the conformation shown, which is arbitrary. Rather the model shows that the fusion protein could fit into the supramolecular structure in a stereochemically acceptable fashion without disturbing the internal structure of either the M13 CP or BPTI domain.

35